Journal of Traumatic Stress, Vol. 20, No. 4, August 2007, pp. 619–623 (© 2007)

Self-Selection Bias in Sleep and Psychophysiological Studies of Posttraumatic Stress Disorder

Steven H. Woodward, Wendy K. Stegman, Joanne R. Pavao, and Ned J. Arsenault
National Center for PTSD, Clinical Laboratory and Education Division, VA Palo Alto HCS, Menlo Park Division, Menlo Park, CA

Tamara L. Hartl
National Center for PTSD, Clinical Laboratory and Education Division, VA Palo Alto HCS, Menlo Park Division, Menlo Park, CA, and Stanford University School of Medicine, Stanford, CA

Kent D. Drescher and Christopher Weaver
National Center for PTSD, Clinical Laboratory and Education Division, VA Palo Alto HCS, Menlo Park Division, Menlo Park, CA

Psychobiological studies of posttraumatic stress disorder (PTSD) often challenge participants to assess the dynamics of systems evolved to organize responses to extreme events. Informed consent insures that volunteers have every opportunity to preevaluate the conditions of the research experience and decline if made uncomfortable by them. Notwithstanding their necessity, these protections set the stage for self-selection phenomena that may bias study outcomes. This study compared prospectively obtained psychometric data from 196 participants and 1229 nonparticipants in sleep and psychophysiological studies of PTSD. Lower subjective nightmare severity was endorsed by persons who later agreed to participate in a study of baseline sleep, an observation consistent with the low nightmare frequencies observed in most laboratories studies of sleep in PTSD.

Many psychobiological studies of posttraumatic stress disorder (PTSD) make significant demands on participants, including exposures to trauma-related cues, lengthy stays in noisy imaging systems, and sleeping in laboratories encumbered by electrodes. It is not surprising then that recruitment for psychobiological studies of PTSD can be difficult. Many potential participants, on learning what awaits them via comprehensive informed consent procedures, decline. These conditions promote the emergence of systematic differences between groups who elect to participate and those who do not.

In this article, we present the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1997) data obtained from 1,425 PTSD inpatients, 196 of whom participated in sleep and/or psychophysiological studies of PTSD, and 1,229 of whom either declined or were never contacted. The data were collected between 1990 and 2000 and comprise a large subsample (∼60%) of all admissions to the inpatient

We wish to acknowledge the contributions of Patricia Chuo, Tracy Thomas, and Gilbert Ramirez who coordinated CAPS data entry over the period of acquisition. We thank Craig Rosen and David Foy for their editorial and statistical advice. We are grateful for the support of the Medical Research Service of the Department of Veterans Affairs and the Research Service of the VA Palo Alto HCS.

Correspondence concerning this article should be addressed to: Steven H. Woodward, PTSD 334, VA Palo Alto HCS, 795 Willow Road., Menlo Park, CA 94025. E-mail: Steve.woodward@med.va.gov.

 DOI: 10.1002/jts.20236

PTSD program associated with the Clinical Laboratory and Education Division of the National Center for PTSD (Menlo Park, CA) over that decade. Participants were male Vietnam combat veterans primarily. Comparisons were made between all 1,229 nonparticipants and those who participated in (a) a laboratory polysomnographic (PSG) study of PTSD-related sleep disturbance under baseline conditions ($n = 120$), (b) a trauma-cue reactivity study involving extensive psychophysiological hook-up and extended audiovisual cue exposures ($n = 49$), and (c) an ambulatory (unattended) PSG study of persons reporting elevated numbers of trauma-related nightmares ($n = 46$). For some comparisons, laboratory study participants were combined into a single group. The null hypothesis was always that no differences existed between participants and nonparticipants.

METHOD

Participants

All participants provided written informed consent following procedures of the Stanford/VA Palo Alto HCS Human Research Protection Program. Laboratory study participants were remunerated at rates typical of VA- and NIMH-funded research. The CAPS interviews were obtained as part of the standard admission process and preceded recruitment into any laboratory study by at least one month. Most were performed by Masters-level practicum students supervised by doctoral staff. Although PTSD assessment accuracy undoubtedly varied, there is no reason to suspect that it varied systematically across groups.

Recruitment of the research participants was mainly accomplished via face-to-face solicitation in the inpatient setting by study staff familiar to the patients. Recruiters had ready access to patients and were highly successful such that approximately two thirds of inpatients contacted agreed to participate in a study in our laboratory. This high success rate probably reflects the strong affiliation between the treatment and research programs and staffs and the fact that inpatients had extra time available for voluntary activities. Based on this rate, we estimate that the nonparticipant sample contained approximately 100 persons who were recruited, but declined, 367 who would have declined if asked, and 733 persons who would have participated if asked. In other words, we estimate that approximately 60% of the nonparticipant group was misclassified.

Data Analysis

After exclusion of all cases with missing or out-of-range values, the final sample consisted of 1,425 interviews. Total severity and criterion scores were normally distributed and analyzed with parametric statistics. All 17 individual symptom severity scores (frequency + intensity) exhibited bimodal distributions characterized by an excess of zeros. Accordingly, individual item severities were recast in a categorical format with scores in the range of 0 to 2 classified as low, 3 to 5 as moderate, and 6 to 8 has high, and analyzed via chi-square analysis. Because we were interested in the possibility that research participation might be associated with variation in the relations among PTSD symptoms, in addition to or instead of differences on individual items, the conventional *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (*DSM-IV*; American Psychiatric Association, 1994) PTSD factor structure was also compared over the main participant and nonparticipant groups using multigroup confirmatory factor analysis (Byrne, 2001). For this model, an overarching PTSD factor led to three separate factors, respectively, representing *DSM-IV* Criterion B—reexperiencing (Factor 1; CAPS items b1–b5), Criterion C—avoidance/numbing (Factor 2; CAPS items c1–c7), and Criterion D—Hyperarousal (Factor 3, CAPS items d1–d5). The model was fitted using maximum likelihood estimation.

How best to control for studywide error rates in this context is unclear as the nonparticipant group included an unknown, but probably large number of persons who would have participated in a laboratory study if contacted. As a result, the observed differences almost certainly underestimate the true differences, biasing the design towards Type II errors. In this context, conventional Type I error control via the Bonferroni method, setting critical alpha to .05/17 or .0029, is probably excessively stringent.

Accordingly, we adopted $p < .05$ as a nominal significance criterion.

RESULTS

Table 1 presents mean CAPS total severity and criterion scores by group. There was some subsample overlap. One person participated in both the laboratory PSG study of baseline sleep and the ambulatory PSG study of trauma-related nightmares. Eighteen participated in both the ambulatory PSG study of trauma-related nightmares and the trauma-cue reactivity study.

Considered as a single group, study participants exhibited slightly lower PTSD severity compared to nonparticipants (78.89 vs. 82.76, respectively), $F(1, 1428) = 6.69$, $p < .01$. A similar result was obtained when only those participants recruited into the laboratory sleep study were compared to nonparticipants (78.90 vs. 82.76, respectively), $F(1, 1347) = 5.05$, $p < .05$. Neither trauma-cue reactivity study participants nor ambulatory nightmare study participants differed in overall PTSD severity from nonparticipants. The MANOVA applied to CAPS B, C, and D criterion scores found no associations between the profile of scores and any of the between-subjects groupings tested: for laboratory sleep versus nonparticipants—Wilks's lambda = 1.00, $F(3, 1345) = 1.92$, *ns*; for trauma cue exposure versus nonparticipants—Wilks's lambda = 1.00, $F(3, 1345) = 1.63$, *ns*; for ambulatory nightmare versus nonparticipants—Wilks's lambda = 1.00, $F(3, 1345) = 1.87$, *ns*. The *DSM-IV* PTSD criterion-based factor structure (factor loadings and variances) did not vary across psychobiological study participants (here considered as a group) and nonparticipants, $\Delta\chi^2(20) = 16.00$, *ns*.

The above tests were followed by a series of planned comparisons of individual CAPS item response distributions across the per-study contrasts (laboratory sleep vs. nonparticipants, trauma cue exposure vs. nonparticipants, ambulatory nightmare vs. nonparticipants). When laboratory sleep study participants ($n = 120$) were compared to nonparticipants, significant differences were observed in endorsements of nightmare complaint, B2: $\chi^2(2, 1349) = 19.75$, $p < .001$; sense of foreshortened future, C7: χ^2 (2, N = 1349) = 11.23, $p < .01$; and irritability, D2: χ^2 (2, N = 1349) = 12.10, $p < .01$. In each case, study participants endorsed lower levels of these three complaints than nonparticipants. (The distributions of CAPS B2 criterion item scores over group are presented in Figure 1.) Despite endorsing lower severities of nightmare complaint, sleep laboratory participants did not endorse lower levels of difficulty falling and staying asleep than nonparticipants, D2: $\chi^2 < 1$.

Compared to nonparticipants, trauma-cue reactivity study participants ($n = 49$) endorsed lower severities on avoidance of activities reminiscent of the trauma, C2: $\chi^2(2, N = 1278) = 6.80$, $p < .05$, and difficulty falling asleep, D1: χ^2 (2, N = 1278) = 8.43, $p < .05$.

Compared to nonparticipants, ambulatory nightmare study participants ($n = 46$) endorsed less distress to cues reminiscent of the trauma, B4: χ^2 (2, 1275) = 11.10, $p < .01$); less avoidance of activities reminiscent of the trauma, D2: χ^2 (2, N = 1275) = 6.46, $p < .05$; less diminishment of interest, C4: χ^2 (2, N = 1275) = 6.26, $p < .05$); and

Table 1. Means and standard deviations (*SD*s) of sleep and psychophysiological study participants and nonparticipants on CAPS summary measures.

	n	CAPS Total	*SD*	B Criterion	*SD*	C Criterion	*SD*	D Criterion	*SD*
Control	1229	82.76	17.90	22.41	7.44	34.19	8.90	26.15	5.80
All lab	196	78.89**	18.24	20.92	7.84	32.53	9.07	25.45	5.69
Lab sleep	120	78.90*	18.31	20.83	7.62	32.98	9.23	25.09	5.71
Waking stressors	49	80.02	20.82	21.66	8.84	32.13	9.51	26.23	6.36
Nightmare complaint	46	79.07	16.19	21.41	6.92	31.48	8.89	26.17	4.99

* $p < .05$. ** $p < .01$.

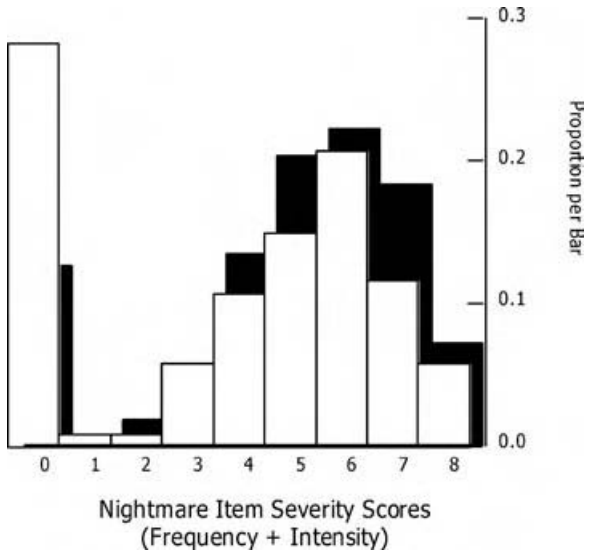

Figure 1. Distributions of CAPS item endorsements for baseline sleep-study participants (white bars) and nonparticipants (black bars). Note a proportional increase in the number of laboratory sleep-study participants denying any nightmare complaints.

less irritability, D2: χ^2 (2, N = 1275) = 8.04, $p < .05$. They did not endorse lower levels of nightmares, per se, B2: χ^2 (2, N = 1275) = .25, *ns*.

DISCUSSION

Considered as a group, sleep and psychophysiological study participants exhibited only slightly lower CAPS total PTSD severity scores than nonparticipants (~3.5 points). There were no associations between any participant/nonparticipant contrast and profiles of B, C, and D criterion scores; and the conventional *DSM* factor structure did not differ between study participants and nonparticipants. These omnibus findings point away from large selection biases and are compatible with a high recruitment success rate.

Examination of item-level's responses told a somewhat different story. Participants in the laboratory sleep study endorsed milder levels of severity on the CAPS sense of foreshortened future item (C7) than nonparticipants. This finding has an appealing interpretation. From the participant's perspective, medical research is an inherently future-oriented activity. Advances resulting from the research will likely benefit persons other than the participant. The second finding is more troubling. It is not clear why participants in a study of baseline sleep should be characterized by lower levels of nightmare complaint rather than higher levels of nightmare complaint or different levels of insomnia complaint. This observation, nevertheless, is wholly compatible with the low rates of nightmares reported in many laboratory studies of sleep in PTSD (Woodward, Arsenault, Murray, & Bliwise, 2000). Furthermore, as participants' reporting histories of trauma-related nightmares may exhibit more objective sleep disturbance in the laboratory than those without (Woodward et al., 2000), a generalized selection bias toward low rates of trauma-related nightmares in PTSD patient samples is compatible with the failure of many laboratory studies to observe reliable sleep architectural modifications (Breslau et al., 2004; Hurwitz, Mahowald, Kuskowski, & Engdahl, 1998; Pillar, Malhotra, & Lavie, 2000). It was also noteworthy that participants endorsing current nightmares, and agreeing to undergo their study, endorsed nightmare severities comparable to the non-participants combined with reduced severity on four other individual symptoms of PTSD.

Additional differences noted between participants and nonparticipants appeared, in a sense, "rational." Participants agreeing to undergo exposure to trauma cues or to report their trauma-related nightmares endorsed lower levels of avoidance of trauma-related memories and stimuli. Participants who volunteered to undergo hours of electrode attachment and detachment over the course of multiple night sleep studies endorsed lower levels of irritability.

This archival study has important limitations. First, it is regrettable that recruitment attempts were not systematically recorded. Misclassification of potential participants to the nonparticipant group can only have operated to reduce differences between participants and nonparticipants; however, we do not know by how much. For this reason as well, conventional alpha protection measures were impossible to implement. Second, these results may not generalize to more common recruitment methods utilizing advertisement and telephone contact. At the same time, insofar

as recruitment rates via those methods are usually much lower than we have enjoyed, the opportunity for selection bias is increased. Because the inpatient milieu provided our participants is explicitly designed to counter the isolative behaviors common among PTSD patients, we do not believe our dependence upon face-to-face recruitment biased our sample towards more socially engaged volunteers.

This study has sampled a limited range of research participant experiences and may simply have stumbled upon an association between PTSD nightmare phenomenology and research participation decision-making. However, it is also possible that similar bias phenomena are at work in other domains of laboratory investigation of PTSD. Unquestionably, some research participants experience neuroimaging, blood draws, and lumbar punctures as challenging. As well, many PTSD patients manifest behavioral tendencies that could bear on recruitability (e.g., avoidance and social isolation). When such factors exercise their influence prior to research participation, they are invisible to researchers. Furthermore, the current data suggest that assessments of selection biases in studies of PTSD should not stop at tests of overall severity or of criterion scores, but also include item analyses when possible. Finally, these findings provide further support for the proposition that trauma-related nightmare symptomatology should be assessed in studies of sleep in PTSD.

REFERENCES

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Charney, D. S., & Keane, T. M. (1997). Clinician-Administered PTSD Scale for DSM-IV: Current and Lifetime Version. Boston: Behavioral Science Division, Boston VA Medical Center/Neurosciences Division, West Haven VA Medical Center.

Breslau, N., Roth, T., Burduvali, E., Kapke, A., Schultz, L., & Roehrs, T. (2004). Sleep in lifetime posttraumatic stress disorder: A community-based polysomnographic study. Archives of General Psychiatry, 61, 508–516.

Byrne, B. M. (2001). Structural equation modeling with AMOS: Basic concepts, applications, & programming. Mahwah, NJ: Erlbaum.

Hurwitz, T. D., Mahowald, M. W., Kuskowski, M., & Engdahl, B. E. (1998). Polysomnographic sleep is not clinically impaired in Vietnam combat veterans with chronic posttraumatic stress disorder. Biological Psychiatry, 44, 1066–1073.

Pillar, G., Malhotra, A., & Lavie, P. (2000). Post-traumatic stress disorder and sleep-what a nightmare! Sleep Medicine Reviews, 4, 183–200.

Woodward, S. H., Arsenault, N. J., Murray, C., & Bliwise, D. L. (2000). Laboratory sleep correlates of nightmare complaint in PTSD inpatients. Biological Psychiatry, 48, 1081–1087.